JOURNAL OF ALGEBRA **89**, 375–396 (1984)

The Nilpotency of the Radical in a Finitely Generated P.I. Ring

AMIRAM BRAUN*

Department of Mathematics, Haifa University, Haifa, Israel

Communicated by I. N. Herstein

Received January 6, 1982

Let $R = \Lambda\{x_1,..., x_k\}$ be a p.i. ring, satisfying a monic polynomial identity (one of its coefficients is ± 1), where Λ is a central noetherian subring. It is proven that $N(R)$, the nil radical of R, is nilpotent. As a corollary, by taking $\Lambda = F$, a field, we settle affirmatively the open problem posed in (C. Procesi, "Rings with Polynomial Identities," p. 186, Marcel Dekker, New York, 1973). We prove: "The Jacobson radical of a finitely generated p.i. algebra is nilpotent."

INTRODUCTION

We shall prove here the following:

THEOREM. *Let $R = \Lambda\{x_1,..., x_k\}$ be a p.i. ring, satisfying a monic polynomial identity, Λ a central noetherian subring.*
Then, $N(R)$, the nil radical of R, is nilpotent.

Here, by a monic polynomial we mean one where some of its coefficients are ± 1.

As a corollary we obtain the next theorem, answering affirmatively the open problem which is proposed in [10, p. 186].

THEOREM. *Let $R = F\{x_1,..., x_k\}$ be a p.i. algebra, where F is a central subfield. Then $N(R)$, the nil radical of R, is nilpotent.*

Kemer has announced a proof of this theorem in the case char $R =$ char $F = 0$. The proof appears in [7]. We summarize his approach. Using the theory of representations of the symmetric group Σ_m over fields of characteristic zero and its connection to p.i. ring theory (e.g., [2, Sect. 6.2]), he is

* A major part of this work was done while the author enjoyed the hospitality of the Institute of Advanced Study, the Hebrew University, Jerusalem, Israel, during the ring theory "workshop," July 1981.

0021-8693/84 $3.00

able to show that *every p.i. algebra of characteristic zero satisfies a Capelli polynomial.* Then he quotes [11, Theorem 3] to complete his proof. This approach does not seem to generalize to algebras of arbitrary characteristics due to the strong dependence of [11, Theorem 3] and the theory of representations of Σ_n on the char $F=0$ assumption.

Using a result of [9, Theorem 10], we obtain the following important corollary.

COROLLARY. *Let $R=F\{x_1,..., x_k\}$ be a p.i. algebra, F a central subfield. Then, R is an homomorphic image of the finitely generated ring of $n \times n$ generic matrices.*

$$F\{\bar{\xi}_1,..., \bar{\xi}_k\}, \qquad \textit{for some } n.$$

The best previously known result concerning the nilpotency of the radical is the following (due in this form to Schelter):

THEOREM [11; 12, pp. 214–215; 14]. *Let $R=\Lambda\{x_1,..., x_k\}$ be a p.i. ring, Λ a central noetherian subring. Suppose that R satisfies all the polynomial identities of $M_n(\mathbb{Z})$, for some n. Then, $N(R)$ is nilpotent.*

The structure of this paper is as follows. Section 0 is devoted to preliminary material. In Section 1 we prove a theorem on T-ideals, generalizing a theorem of Razmyslov [11, Theorem 3]. Perhaps this section should be skipped in the first reading (although it is used in Section 2). The main results are proved in Section 2. Here we first prove the main theorem with the additional assumption Krull-dim$(\Lambda) < \infty$. This is done in Theorem 2.2. The affirmative solution to the open problem in [10, pp. 103, 186] already follows. The main theorem is proved in Theorem 2.5. This separation is done in order to clarify the inductive procedure which is much more easier in Theorem 2.2. In fact this is the sole difference between the two theorems. Finally, in the Appendix we reproduce a result of Latyshev, which we heavily use, together with a theorem of Shirshov on which it depends.

0. PRELIMINARIES

We shall use the following notations and conventions throughout this paper. $Z(R)$ will denote the center of R and $N(R)$ will denote the nil radical of R.

Throughout this work $R=\Lambda\{x_1,..., x_k\}$ will denote an associative, finitely generated p.i. ring, with 1, where Λ is a central noetherian subring. Then by [10, p. 108, Corollary 2.4], $N(R)=P_1 \cap \cdots \cap P_t$, where $\{P_i\}$ are the minimal prime ideals of R.

An instrumental tool in this investigation is the following:

THEOREM 0.1 [8, Proposition 12]. *Let $R = A\{x_1,..., x_k\}$ be a p.i. ring satisfying a monic polynomial identity. Let $I \subset N(R)$ be a two-sided ideal. Then, I is nilpotent provided I is finitely generated as a two-sided ideal.*

Remark 0.2. (1) The statement of Proposition 12 in [8] involves the assumption that $A = F$ is a field and $\operatorname{char} F = 0$. These *restrictions* are obviously *superfluous*. Indeed, Shirshov's theorem [16, Theorem 2], which is the only external tool to be used in [8, Proposition 12], is combinatorial in nature and is "characteristic free" (e.g., [12, p. 206]). (See Appendix for a detailed explanation.)

(2) By a monic polynomial we mean a polynomial at least one of whose coefficients is ± 1. This is the only restriction in Shirshov's theorem [16, Theorem 2; 12, p. 206].

We use the notations $R_\lambda \equiv R[1/\lambda]$ to denote the localization of R by $\lambda \in Z(R)$ (e.g., [12, p. 51]). Observe that λ may be a zero divisor in R. In fact if $\nu: R \to R_\lambda$ denotes the obvious map $r \to r \cdot 1^{-1}$ then $\ker \nu = \{x \in R \mid \lambda^s x = 0, \text{ for some } s\}$.

We shall have cause to use $R^e \equiv R \otimes_z R^0$, the enveloping algebra of R, where $Z \equiv Z(R)$. One has the following natural map:

$$\nu_\lambda: R \otimes_z R^0 \to R_\lambda \otimes_{z_\lambda} R_\lambda^0 = (R_\lambda)^e = (R^e)_\lambda,$$

defined by $\nu_\lambda(\sum a_i \otimes_z b_i) = \sum a_i \otimes_{z_\lambda} b_i$. We have that

$$\ker \nu_\lambda = \{x \in R^e \mid \lambda^s x = 0 \text{ for some } s\}.$$

We also have a map ϕ: $R^e \to \operatorname{Hom}_z(R, R)$, defined by $\{\phi(\sum a_i \otimes_z b_i)\}(x) = \sum a_i x b_i$, for all $x \in R$. It is easily verified that ϕ is a ring homomorphism.

$(0_r: J(R))$ denotes $\{d \in R^e \mid (x \otimes 1 - 1 \otimes x^0)d = 0, \text{ for every } x \text{ in } R\}$.

One has the following:

THEOREM 0.3. (1) *If R is an Azumaya algebra, then $R^e = R^e(0_r: J(R))$.*

(2) *Let $d \in (0_r: J(R))$, $d = \sum a_i \otimes b_i^0$. Then $\sum a_i x b_i \in Z(R)$, for every $x \in R$.*

(3) *If $d \in R^e$, $R = A\{x_1,..., x_k\}$ and $(x_i \otimes 1 - 1 \otimes x_i^0)d = 0$ for $i = 1,..., k$. Then, $d \in (0_r: J(R))$.*

Proof. We prove (1). By [6, Theorem 3.4, p. 52], R being Azumaya implies that R is a generator as a left R^e module. Now, by [6, p. 53, line 10] this is equivalent to the fact that $R^e(0_r: J(R)) = R^e$.

To prove (2) we use $\phi: R^e \to \operatorname{Hom}_z(R, R)$. $\sum a_i \otimes b_i^0 \in (0_r : J(R))$, hence $(r \otimes 1 - 1 \otimes r^0)(\sum a_i \otimes b_i^0) = 0$ for all $r \in R$; equivalently $\sum ra_i \otimes b_i^0 = \sum a_i \otimes (b_i r)^0$, for all $r \in R$. Apply ϕ to this equality and get $\{\phi(\sum ra_i \otimes b_i^0)\}(x) = \{\phi(\sum a_i \otimes (b_i r)^0)\}(x)$, for all $x \in R$; consequently $r(\sum a_i x b_i) = (\sum a_i x b_i) r$ for all $x, r \in R$.

The proof of (3) appears in [4, Proposition 2.1]. We repeat the reasoning for the sake of completeness. We shall show that $[(xy \otimes 1)d = (1 \otimes (xy)^0)d$ provided $(x \otimes 1)d = (1 \otimes x^0)d$, $(y \otimes 1)d = (1 \otimes y^0)d$. Indeed $(xy \otimes 1)d = (x \otimes 1)(y \otimes 1)d = (x \otimes 1)(1 \otimes y^0)d = (1 \otimes y^0)(x \otimes 1)d = (1 \otimes y^0)(1 \otimes x^0)d = (1 \otimes (xy)^0)d]$.

Now, one argues by induction on the length of monomials in $x_1, \dots, x_k$ that $(r \otimes 1 - 1 \otimes r^0)d = 0$ for each such monomial r (the first step of the induction is $(x_i \otimes 1 - 1 \otimes x_i^0)d = 0$ for $i = 1, \dots, k$). Now by linearity and the fact that $\Lambda \subseteq Z(R)$ we get that $(r \otimes 1 - 1 \otimes r^0)d = 0$ for all $r \in R$.

The next lemma is part of the folklore of p.i. ring theory.

LEMMA 0.4. *Let $R = \Lambda\{x_1, \dots, x_k\}$ be a p.i. ring (Λ a central noetherian subring). Suppose that R/I is a finite module over its center, and I is a two-sided ideal in R.*

Then I is a finitely generated two-sided ideal.

Proof. Let $\bar{u}_1, \dots, \bar{u}_m$ be the generators of R/I over $Z(R/I)$. Then

$$\bar{u}_i \bar{u}_j = \sum_{t=1}^{m} \bar{\alpha}_{ijt} \bar{u}_t, \qquad \bar{\alpha}_{ijt} \in Z\left(\frac{R}{I}\right), \qquad t, i, j = 1, \dots, m,$$

$$\bar{x}_i = \sum_{j=1}^{m} \bar{\beta}_{ij} \bar{u}_j, \qquad \bar{\beta}_{ij} \in Z\left(\frac{R}{I}\right), \qquad i = 1, \dots, k, j = 1, \dots, m.$$

Let I_0 be the two-sided ideal in R generated by

$$u_i u_j - \sum_t \alpha_{ijt} u_t, \ [\alpha_{ijt}, x_l], \qquad t, i, j = 1, \dots, m, l = 1, \dots, k,$$

and by

$$x_i - \sum_j \beta_{ij} u_j, \ [\beta_{ij}, x_l], \qquad i, l = 1, \dots, k, j = 1, \dots, m.$$

Then $I_0 \subset I$, I_0 is a finitely generated two-dsided ideal in R. Also, $\tilde{R} \equiv R/I_0$ is a finite module over its central subring $\Lambda[\tilde{\alpha}_{ijt}, \tilde{\beta}_{rj}, \ i, j, t = 1, \dots, m, r = 1, \dots, k]$. Thus $\tilde{R}$ is left and right noetherian and I/I_0 is a finitely generated two-sided ideal. Consequently, I is finitely generated.

COROLLARY 0.5. *Let $R = \Lambda\{x_1, \dots, x_k\}$ be a p.i. ring satisfying a monic polynomial identity, Λ a central noetherian subring.*

Let I be a two-sided ideal such that

(i) *R/I is a finite module over $Z(R/I)$;*

(ii) *$I \subseteq N(R)$.*

Then, I is a nilpotent ideal.

Proof. We apply Theorem 0.1 and Lemma 0.4. Q.E.D.

We now make two additional definitions. Firstly, by p.i.d.(R) we denote the minimal number n (if it exists), such that R satisfies all the identities of $n \times n$ matrices $(M_n(\mathbb{Z}))$ over $\mathbb{Z}$. We also write p.i.d.$(\{0\}) = 0$.

Second, by k.d.(R), we denote the classical Krull-dimension of R, that is, the maximal length of chains of prime ideals of R. One has k.d.$(R) =$ k.d.$(R/N(R))$.

LEMMA 0.6. *Let $R = \Lambda\{x_1, ..., x_k\}$ be a p.i. ring, Λ a central noetherian subring. Suppose that* k.d.$(\Lambda) < \infty$. *Then* k.d.$(R) \leqslant b < \infty$, *where $b = b(d, k, \Lambda)$ is a function of d, k, Λ, and d is the degree of the identity which R satisfies.*

Proof. The proof, which uses [10, p. 179] or [12, Exercise 1.10.1, p. 106] together with [10, Proposition 2.3, p. 107], is left to the reader.

We shall use the notation $[a, b] = ab - ba$, $a, b \in R$. Finally, let $f(\bar{\mathbf{X}}_1, ..., \bar{\mathbf{X}}_{d^2}, \bar{\mathbf{Y}}_1, ..., \bar{\mathbf{Y}}_s)$ be the multiplinear, central polynomial, alternating in $\bar{\mathbf{X}}_1, ..., \bar{\mathbf{X}}_{d^2}$, obtained from the Capelli polynomial C_{d^2}, by S. Amitsur and G. Bergman from Razmyslov's central polynomial; cf. [12, pp. 26–27].

The following generalization of the Artin–Procesi theorem was obtained by W. Schelter.

THEOREM 0.7 (1, p. 10;12, Theorem 1.8.33; 14). *Let R be a p.i. ring, $1 \in R$, satisfying the identity $[f(\bar{\mathbf{X}}_1, ..., \bar{\mathbf{X}}_{d^2}, \bar{\mathbf{Y}}_1, ..., \bar{\mathbf{Y}}_s), \bar{\mathbf{X}}] = 0$. Suppose that* p.i.d.$(R/M) = d$ *for every maximal ideal M in R. Then R is an Azumaya algebra of constant rank d^2.*

ASSUMPTION 0.8. (1) *We may and shall assume without further notice that $P(X_1, ..., X_d)$, the identity which R satisfies, is multilinear in $X_1, ..., X_d$.*

(2) *We may further assume that $\exists \eta \in \Lambda$ so that $\prod_{1 \leqslant i < j \leqslant d} (\eta^i - \eta^i) \equiv \xi$ is invertible in R. In fact, adding a new variable t and taking $\eta = t$, one verifies that $R \subseteq R[t] \subseteq R[t, 1/\xi]$, since ξ is regular in $R[t]$, and it suffices to prove the theorem for the latter ring.*

1. A Generalization of a Theorem of Razmyslov

Let F_m be the free algebra generated by $a_1,..., a_k$, $x_1,..., x_m$ over Λ. Let $\Lambda\{a_1,..., a_k\} = A$, be the free subalgebra of F_m, generated by $a_1,..., a_k$, and $Z \equiv Z(A)$, the center of A. Recall the mth Capelli polynomial:

$$d_m(x_1,..., x_m, y_0,..., y_m) = \sum_{\sigma \in \Sigma_m} (sg\sigma)\, y_0 x_{\sigma(1)} \cdots y_{m-1} x_{\sigma(m)} y_m.$$

We add to $A \otimes_Z 1 \subseteq A \otimes_Z A^0$ new centralizing indeterminates $\{t_i(h) \mid i = 1,..., m, h \in A\}$. Let $B = (A \otimes_Z 1)\{t_i(h) \mid i = 1,..., m, h \in A\}$. Let $L = \mathrm{span}_\Lambda\{d_m(x_1,..., x_m, y_0,..., y_m) \mid y_j \in A$, is a monomial, $j = 0,..., m\}$.

We now regard $L \subseteq F_m$ as a left B module via the following action:

(i) $(a \otimes 1) \cdot d_m(x_1,..., x_m, y_0,..., y_m) = a d_m(x_1,..., x_m, y_0,..., y_m)$, where $a \in A$;

(ii) $t_l(h) \cdot d_m(x_1,..., x_m, y_0,..., y_m) = \sum_{1 \leqslant i_1 < \cdots < i_l \leqslant m} d_m(x_1,..., hx_{i_1},..., hx_{i_l},..., x_m, y_0,..., y_m)$, where $h \in A$, $l = 1,..., m$.

This last definition is best motivated by adding a new variable λ to F_m, and realizing that (ii) is obtained by computing the coefficient of λ^{m-l} in the following equality:

(ii′) $(\lambda^m - t_1(h)\lambda^{m-1} + \cdots \pm t_m(h)) \cdot d_m(x_1,..., x_m, y_0,..., y_m) = d_m((\lambda - h)x_1,..., (\lambda - h)x_m, y_0,..., y_m)$,

which is actually equivalent to (ii).

We next prove the following identity.

(iii) $t_l(h)\, d_m(x_1,..., x_m, y_0,... y_m) = \sum_{1 \leqslant k_1 < k_2 \cdots < k_l \leqslant m} d_m(x_1,..., x_m, y_0,..., (y_{k_1 - 1}h), y_{k_1},..., (y_{k_l - 1}h),..., y_{k_l},..., y_m)$.

Indeed, let $i_j \in \{1,..., m\}$, $j = 1,..., l$, and $i_j \neq i_s$ if $j \neq s$ and let $k_j \in \{1,..., m\}$, $j = 1,..., l$ and $k_j \neq k_s$ if $j \neq s$. Let

$$B(k_1, i_1, k_2, i_2,..., k_l, i_l) = \{\sigma \in \Sigma_m \mid \sigma(k_1) = i_1,..., \sigma(k_l) = i_l\}.$$

Then one easily verifies that

$$\Sigma_m = \dot{\bigcup_{1 \leqslant k_1 < \cdots < k_l \leqslant m}} B(k_1, i_1,..., k_l, i_l) = \dot{\bigcup_{1 \leqslant i_1 < \cdots < i_l \leqslant m}} B(k_1, i_1,..., k_l, i_l),$$

where $\dot{\bigcup}$ denotes a disjoint union.

We have

$$d_m(x_1,\dots, hx_{i_1},\dots, hx_{i_l},\dots, x_m, y_0,\dots, y_m)$$
$$= \sum_{\sigma\in\Sigma_m} (sg\sigma)\, y_0 x_{\sigma(1)} \cdots y_{\sigma^{-1}(i_1)-1}(hx_{i_1})\, y_{\sigma^{-1}(i_1)}$$
$$\cdots y_{\sigma^{-1}(i_l)-1}(hx_{i_l})\, y_{\sigma^{-1}(i_l)} \cdots x_{\sigma(m)}\, y_m.$$

Let $\sigma^{-1}(i_1) = k_1,\dots, \sigma^{-1}(i_l) = k_l$; then

$$d_m(x_1,\dots, hx_{i_1},\dots, hx_{i_l},\dots, x_m, y_0,\dots, y_m)$$
$$= \sum_{\sigma\in\Sigma_m} (sg\sigma)\, y_0 x_{\sigma(1)} \cdots (y_{k_1-1}h)\, x_{\sigma(k_1)}\, y_{k_1}$$
$$\cdots (y_{k_l-1}h)\, x_{\sigma(k_l)}\, y_{k_l} \cdots x_{\sigma(m)}\, y_m$$
$$= \sum_{1\leqslant k_1<\cdots<k_l\leqslant m}\ \sum_{\sigma\in B(k_1,i_1,\dots,k_l,i_l)} (sg\sigma)\, y_0 x_{\sigma(1)} \cdots (y_{k_1-1}h)\, x_{\sigma(k_1)}$$
$$\cdots (y_{k_l-1}h)\, x_{\sigma(k_l)}\, y_{k_l} \cdots x_{\sigma(m)}\, y_m.$$

Consequently,

$$t_l(h)\, d_m(x_1,\dots, x_m, y_0,\dots, y_m)$$
$$= \sum_{1\leqslant i_1<\cdots<i_l\leqslant m} d_m(x_1,\dots, hx_{i_1},\dots, hx_{i_l},\dots, x_m, y_0,\dots, y_m)$$
$$= \sum_{1\leqslant i_1<\cdots<i_l} \Bigg(\sum_{1\leqslant k_1<\cdots<k_l\leqslant m}\ \sum_{\sigma\in B(k_1,i_1,\dots,k_l,i_l)} (sg\sigma)\, y_0 x_{\sigma(1)}$$
$$\cdots (y_{k_1-1}h)\, x_{\sigma(k_1)} \cdots (y_{k_l-1}h)\, x_{\sigma(k_l)} \cdots x_{\sigma(m)}\, y_m$$
$$= \sum_{1\leqslant k_1<\cdots<k_l\leqslant m} \Bigg(\sum_{\sigma\in\Sigma_m} (sg\sigma)\, y_0 x_{\sigma(1)} \cdots (y_{k_1-1}h)\, x_{\sigma(k_1)}$$
$$\cdots (y_{k_l-1}h)\, x_{\sigma(k_l)} \cdots x_{\sigma(m)}\, y_m \Bigg)$$
$$= \sum_{1\leqslant k_1<\cdots<k_l\leqslant m} d_m(x_1,\dots, x_m, y_0,\dots, (y_{k_1-1}h),\dots, (y_{k_l-1}h),\dots, y_m).$$

Q.E.D.

We have now the following:

LEMMA 1.1. *Let* $L = \mathrm{Span}_\Lambda\{d_m(x_1,\dots, x_m, y_0,\dots, y_m) \mid y_j \in A,$ *is a monomial*, $j = 0,\dots, m\}$.

Then, L *is a left* B *module*.

Proof. Clearly, by (i), (iii), and (ii),

$$\{(a \otimes 1)\, t_l(h) - t_l(h)(a \otimes 1)\} \cdot d_m = 0,\ a, h \in A;$$

then one shows that L is closed under the action of the generators of $B = (A \otimes_Z 1)\{t_i(h) \mid h \in A,\ i = 1,..., m\}$.

Remarks. (1) If $l = 1$, $m = \pi^2$, then

$$t_1(h)\, d_m = \sum_{i=1}^{m} d_m(x_1,..., hx_i,..., x_m, y_0,..., y_m)$$

which already appears in [11, p. 201].

(2) For the motivation of (ii) and (ii′) with $m = \pi^2$, we refer to [1, (2.4), (2.5)].

We assume now that $m = \pi^2$ and by using (ii′), one readily checks that

$$\text{(iv)}\qquad \left\{ t_{\pi^2}(\lambda - h) - \sum_{i=0}^{\pi^2} \lambda^i t_{\pi^2 - i}(h)(-1)^{\pi^2 - i} \right\} \cdot d_{\pi^2}(x_1,..., x_{\pi^2},..., y_{\pi^2}) = 0;$$

consequently

$$\left\{ t_{\pi^2}(\lambda - h) - \sum_{i=0}^{\pi^2} \lambda^i t_{\pi^2 - i}(h)(-1)^{\pi^2 - i} \right\} \cdot L = \{0\}, \qquad \text{for} \quad h \in A, \lambda \in \Lambda.$$

Let $U = (u_{ij})$ be a $(\pi^2 + 1) \times (\pi^2 + 1)$ matrix, where $u_{ij} = (\eta^i)^{\pi^2 - j + 1}$, and η is taken as in Assumption 0.8. Then by (ii′) we have

$$U \cdot \begin{pmatrix} (-1)t_1^1(h) \\ \vdots \\ (-1)^{\pi^2} t_{\pi^2}(h) \end{pmatrix} \cdot d_{\pi^2}(x_1,..., y_{\pi^2}) = \begin{pmatrix} t_{\pi^2}(\eta - h) \\ \vdots \\ t_{\pi^2}(\eta^{\pi^2+1} - h) \end{pmatrix} \cdot d_{\pi^2}(x_1,..., y_{\pi^2}).$$

Now, since U, by the Vandermandé argument, is a nonsingular matrix, we may multiply both sides of the equality by U^{-1} and deduce that

$$\left\{ t_j(h) - \sum_{i=1}^{\pi^2+1} \alpha_{ij} t_{\pi^2}(\eta^i - h) \right\} \cdot d_{\pi^2}(x_1,..., x_{\pi^2}, y_0,..., y_{\pi^2}) = 0,$$

$$j = 1,..., \pi^2, \text{ where } \{\alpha_{ij}\} \subseteq \Lambda.$$

Consequently,

$$\begin{aligned} &[[t_{i_1}(h_1), t_{i_2}(h_2)],..., t_{i_r}(h_r)] \cdot L \\ &\quad \subseteq \sum [[t_{\pi^2}(\eta^{j_1} - h_1), t_{\pi^2}(\eta^{j_2} - h_2)],..., t_{\pi^2}(\eta^{j_r} - h_r)] \cdot L \\ &\quad \subseteq \sum_{\alpha, \beta, a, b} [t_{\pi^2}(\alpha - a), t_{\pi^2}(\beta - b)] \cdot L, \qquad \text{where} \quad \alpha, \beta \in \Lambda,\ a, b \in A. \end{aligned}$$

The last inclusion holds since $t_{\pi^2}(x)\, t_{\pi^2}(y)\, d_{\pi^2}(x_1,..., x_{\pi^2}, y_0,..., y_{\pi^2}) = t_{\pi^2}(yx)\, d_{\pi^2}(x_1,..., x_{\pi^2}, y_0,..., y_{\pi^2})$, and therefore $[t_{\pi^2}(a), t_{\pi^2}(b), t_{\pi^2}(c)]\, d_{\pi^2} = [t_{\pi^2}(ba) - t_{\pi^2}(ab), t_{\pi^2}(c)]\, d_{\pi^2} = [t_{\pi^2}(ba), t_{\pi^2}(c)]\, d_{\pi^2} - [t_{\pi^2}(ab), t_{\pi^2}(c)]\, d_{\pi^2}$, and similarly for longer commutators. Let

$$\begin{aligned} f_{\pi^2} &\equiv f_{\pi^2}(Y, X, X_1,..., X_{\pi^2}, Y_0,..., Y_{\pi^2}) \\ &= [t_{\pi^2}(X), t_{\pi^2}(Y)] \cdot d_{\pi^2}(X_1,..., X_{\pi^2}, Y_0,..., Y_{\pi^2}); \end{aligned}$$

then obviously, f_{π^2} is a polynomial identity of $M_\pi(\mathbb{Z})$ and

$$[...[t_{i_1}(h_1), t_{i_2}(h_2)],..., t_{i_r}(h_r)] \cdot L \subseteq V(f_{\pi^2}), \tag{v}$$

where $V(f_{\pi^2})$ is the ideal in F_{π^2} generated by all the evaluations of f_{π^2} (where $X, Y, X_1,..., Y_{\pi^2}$ are in A). We are now able to prove, using our previous notations:

THEOREM 1.2. *$V_c^t(A) \cdot L \subset V$ for some c and t, where $V_c(A)$ denotes the ideal in A generated by all the evaluations of the identities of $M_c(\mathbb{Z})$ and V denotes the ideal generated in F_{π^2} by all the evaluations of the identities f_{π^2} and*

$$\begin{aligned} g &= g(X, X_1,..., X_{\pi^2}, Y_0,..., Y_{\pi^2}) \\ &\equiv (X^{\pi^2} - t_1(X)\, X^{\pi^2-1} + t_2(X)\, X^{\pi^2-2} - \cdots \pm t_{\pi^2}(X)) \\ &\quad \times d_{\pi^2}(X_1,..., X_{\pi^2}, Y_0,..., Y_{\pi^2}). \end{aligned}$$

Each one of these identities is a polynomial identity of $M_\pi(\mathbb{Z})$.

Proof. Let J be the ideal in $B = (A \otimes 1)\{t_i(h) \mid h \in A,\ i = 1,..., \pi^2\}$ generated by $\{[t_i(h), t_j(h')] \mid i, j = 1,..., \pi^2, h, h' \in A\}$ and by $\{(h \otimes 1)^{\pi^2} - t_1(h)(h \otimes 1)^{\pi^2-1} + \cdots \pm t_{\pi^2}(h) \mid h \in A\}$. We claim that $J \cdot L \subseteq V$. Indeed, $[t_i(h), t_j(h')]\, d_{\pi^2} \in V(f_{\pi^2}) \subseteq V$ by (v) and $\{(h \otimes 1)^{\pi^2} - t_1(h)(h \otimes 1)^{\pi^2+1} + \cdots \pm t_{\pi^2}(h)\}\, d_{\pi^2} \in V(g) \subseteq V$ by definition. Now, if α is a generator of J then $t_l(z)\, \alpha d_{\pi^2} = [t_l(z), \alpha]\, d_{\pi^2} + \alpha t_l(z)\, d_{\pi^2}$ and $\alpha t_l(z)\, d_{\pi^2} \in \alpha L \subseteq V$ by the previous observations. Also, $[t_l(z), \alpha]\, d_{\pi^2} \in V(f_{\pi^2})$ by (iv) and (v). Similarly, we show that $t_{l_1}(z_1) \cdots t_{l_r}(z_r)\, \alpha d_{\pi^2} \in V$. Indeed, one easily checks that $t_{l_1}(z_1) \cdots t_{l_r}(z_r)\, \alpha d_{\pi^2} \in \sum_{j=0}^{r} [t_{i_1}(z_1), [t_{i_2}(z_2),..., [t_{i_j}, \alpha]] \cdot L \subseteq V(f_{\pi^2}) \subseteq V$, where the second inclusion is obtained by (v) and the first inclusion is due to Lemma 1.1 and repeated commutations. Consequently, $J \cdot L \subseteq V$.

Now, in $B/J \equiv \bar{B}$, $\bar{t}_i(h)$ is central for $i = 1,..., \pi^2$, $h \in A$, and each element $\overline{h \otimes 1} \in \overline{A \otimes 1}$ is integral over $z(\bar{B})$ with bounded degree π^2. Consequently, $\overline{A \otimes 1}$ satisfies a p.i. of degree $1 + 2 + \cdots + (\pi^2 + 1) = e$ (e.g., [10, p. 18]). Let $B' = (A \otimes 1)\{t_i(h) \mid i = 1,..., \pi^2,\ h$ is a monomial of length $\leqslant e\}$; then $B'/B' \cap J$ is a finite module over its noetherian center by Shirshov's theorem

[12, p. 206], and so its nil radical is nilpotent. Now, since $V_c(B') \subseteq N(B'/B' \cap J)$ for some c, there exists t such that $V_c^t(B') \subseteq J$, and therefore $V_c^t(A \otimes 1) \subseteq J$. Consequently,

$$V_c^t(A) \cdot L = V_c^t(A \otimes 1) \cdot L \subseteq J \cdot L \subseteq V.$$

COROLLARY 1.3. *Let $R = A\{a_1', ..., a_k'\}$ be a ring ($A \subset Z(R), \exists \eta \in A$, $\prod_{i<j} (\eta^i - \eta^j)$ is invertible) satisfying the identities f_{π^2}, g (of $M_\pi(\mathbb{Z})$). Then, there exists c such that $V_c^t(R)\, V(d_{\pi^2}) = \{0\}$ for some t, where $V_c(R)$ is the ideal in R, generated by all the evaluations of the identities of $M_c(\mathbb{Z})$, and $V(d_{\pi^2})$ is the ideal in R generated by all the evaluations of d_{π^2}.*

Proof. There exists an epimorphism φ from F_{π^2} onto R, so, by Theorem 1.2 we get $V_c^t(\varphi(A)) \cdot \varphi(L) \subseteq \varphi(V) = \{0\}$. Now we use $\varphi(L) = V(d_{\pi^2})$, $V_c(\varphi(A)) = V_c(R)$.

2. PROOF OF MAIN RESULTS

PROPOSITION 2.1. *Let $R = A\{x_1, ..., x_k\}$ be a ring, $\lambda \in Z(R)$ so that $R_\lambda \equiv R[1/\lambda]$ is Azumaya. Then there exist $b_1, ..., b_h \in R$ such that*

(i) *$Zb_1 + \cdots Zb_h$ is a ring;*

(ii) *$\lambda^e R \subseteq Zb_1 + \cdots + Zb_h$, for some e.*

Proof. By Theorem 0.3(1) we have $R_\lambda^e = R_\lambda^e(0_r : J(R_\lambda))$. Consequently,

$$1 \otimes_{z_\lambda} 1 = \sum_{i=1}^{h} \left(\sum_j c_{ij} \otimes_{z_\lambda} d_{ij}^0\right)\left(\sum_j v_{ij} \otimes_{z_\lambda} w_{ij}^0\right),$$

where $\sum_j v_{ij} \otimes_{z_\lambda} w_{ij}^0 \in (0_r : J(R_\lambda))$, for $i = 1, ..., h$. Multiplying by λ^{e_1} for some e_1, we have that $c_{ij}' \equiv \lambda^{e_1} c_{ij}$, $\lambda^{e_1} d_{ij} \equiv d_{ij}'$, $\lambda^{e_1} v_{ij} \equiv v_{ij}'$, $\lambda^{e_1} w_{ij} \equiv w_{ij}'$ are in R. Thus,

$$\lambda^{4e_1} \otimes_{z_\lambda} 1 = \sum_{i=1}^{h} \left(\sum_j c_{ij}' \otimes_{z_\lambda} d_{ij}'^0\right)\left(\sum_j v_{ij}' \otimes_{z_\lambda} w_{ij}'^0\right),$$

where

$$\sum_j v_{ij}' \otimes_{z_\lambda} w_{ij}'^0 = \lambda^{2e_1}\left(\sum_j v_{ij} \otimes_{z_\lambda} w_{ij}^0\right) \in (0_r : J(R_\lambda))$$

$$\text{for} \quad i = 1, ..., h. \quad (1)$$

Let $\nu_\lambda : R \otimes_z R^0 \to R_\lambda \otimes_{z_\lambda} R_\lambda^0$ be the natural map, $\nu_\lambda(\sum a_j \otimes_z b_j^0) \equiv \sum a_j \otimes_{z_\lambda} b_j^0$. Then if $X \in \ker \nu_\lambda$, there exists an s such that $\lambda^s \cdot X = 0$.

Now, by the very definition of $(0_r : J(R_\lambda))$,

$$(x_l \otimes_{z_\lambda} 1 - 1 \otimes_{z_\lambda} x_l^0) \left(\sum_j v'_{ij} \otimes_{z_\lambda} w'^0_{ij}\right) = 0$$

for $i = 1,..., h$, $l = 1,..., k$. Consequently

$$\nu_\lambda \left\{(x_l \otimes_z 1 - 1 \otimes_z x_l^0) \left(\sum_j v'_{ij} \otimes_z w'^0_{ij}\right)\right\} = 0, \qquad i = 1,..., h,\ l = 1,..., k.$$

Let λ^{e_2} be such that

$$\lambda^{e_2} \left\{(x_l \otimes_z 1 - 1 \otimes_z x_l^0) \left(\sum_j v'_{ij} \otimes_z w'^0_{ij}\right)\right\} = 0 \qquad i = 1,..., h,\ l = 1,..., k.$$

Then, if $f_i(\ ) \equiv \sum_j \lambda^{e_2} v'_{ij} \otimes_z w'^0_{ij}$, $i = 1,..., h$, we have by Theorem 0.3(3) that $f_i(\ ) \in (0_r : J(R))$, $i = 1,..., h$. Now by (1) we have that

$$B \equiv \left\{\lambda^{4e_1} \otimes_z 1 - \sum_{i=1}^{h} \left(\sum c'_{ij} \otimes_z d'^0_{ij}\right)\left(\sum_j v'_{ij} \otimes w'^0_{ij}\right)\right\} \in \ker \nu_\lambda .$$

Let e_3 be such that $\lambda^{e_3} \cdot B = 0$. Then, setting $e_0 = 4e_1 + e_2 + e_3$ one has that

$$\lambda^{e_0} \otimes_z 1 = \sum_{i=1}^{h} \left(\sum \lambda^{e_3} c'_{ij} \otimes_z d'^0_{ij}\right) f_i(\ ).$$

Let ϕ be the natural map $\phi : R \otimes_z R^0 \to \mathrm{Hom}_z(R, R)$, defined by $\{\phi(a \otimes b^0)\}(r) \equiv arb$ and extend linearly. Then $\{\phi(f_i(\ ))\}(r) = \sum_j \lambda^{e_2} v'_{ij} r w'_{ij} \equiv f_i(r) \in Z(R)$ for every $r \in R$ by Theorem 0.3(2). Consequently,

$$\begin{aligned}
\lambda^{e_0} r = \{\phi(\lambda^{e_0} \otimes_z 1)\}(r) &= \sum_{i=1}^{h} \left\{\sum_{i=1}^{h} \left\{\phi\left(\sum_j \lambda^{e_3} c'_{ij} \otimes_z d'_{ij}\right) \cdot \phi(f_i(\ ))\right\}(r)\right. \\
&= \sum_{i=1}^{h} \left\{\phi\left(\sum_j \lambda^{e_3} c'_{ij} \otimes_z d'^0_{ij}\right)\right\}(f_i(r)) \\
&= \sum_{i=1}^{h} \left(\sum_j \lambda^{e_3} c'_{ij} f_i(r)\, d'_{ij}\right) = \sum_{i=1}^{h} \left(\sum_j \lambda^{e_3} c'_{ij} d'_{ij}\right) f_i(r) \\
&= \sum_{i=1}^{h} f_i(r)\, \varepsilon_i, \qquad \text{where} \quad \varepsilon_i = \sum_j \lambda^{e_3} c'_{ij} d'_{ij}, \quad i = 1,..., h
\end{aligned}$$

since $f_i(r) \in Z(R)$, $i = 1,\ldots, h$, for every $r \in R$. Let $b_i \equiv \lambda^{e_0}\varepsilon_i$, $i = 1,\ldots, h$. Then

$$b_i b_j = \lambda^{2e_0}\varepsilon_i\varepsilon_j = \lambda^{e_0}(\lambda^{e_0}\varepsilon_i\varepsilon_j) \in \lambda^{e_0}(\lambda^{e_0}R) \subseteq \lambda^{e_0}\left(\sum_{i=1}^{h} f_i(R)\,\varepsilon_i\right)$$
$$= \sum_{i=1}^{h} f_i(R)\,b_i \subseteq \sum_{i=1}^{h} Z(R)\,b_i.$$

Consequently, if $e \equiv 2e_0$, $\lambda^e R \subseteq Zb_1 + \cdots + Zb_h$, where $Z = Z(R)$, $b_1,\ldots, b_h \in R$ and $Zb_1 + \cdots + Zb_h$ is a ring.

We prove the following:

THEOREM 2.2. *Let $R = \Lambda\{x_1,\ldots, x_k\}$ be a* p.i. *ring, satisfying a monic polynomial identity, Λ a central noetherian subring such that $k \cdot d(\Lambda) < \infty$. Then $N(R)$ is nilpotent.*

Proof. Let $N(R) = P_1 \cap \cdots \cap P_t$, where $\{P_i\}$ are the minimal prime ideals in R, ordered such that

$$\text{p.i.d.}\left(\frac{R}{P_1}\right) = \cdots = \text{p.i.d.}\left(\frac{R}{P_m}\right) \gneqq \text{p.i.d.}\left(\frac{R}{P_{m+1}}\right) \geqslant \cdots \geqslant \text{p.i.d.}\left(\frac{R}{P_t}\right)$$
$$\text{for some } m \leqslant t (P_{m+1} \equiv R \text{ if } m = t).$$

Let $\pi(R) \equiv \text{p.i.d.}(R/P_1)$, $d(R) = \max\{\text{k.d.}(R/P_i) \mid 1 \leqslant i \leqslant m\}$.

We shall restrict attention to the following set:

$$\Gamma = \{\langle \pi(S), d(S)\rangle \mid S = \Lambda\{y_1,\ldots, y_k\},\ 1 \in S,\ S \text{ is a ring}$$
$$\text{satisfying } P(\bar{\mathbf{X}}_1,\ldots, \bar{\mathbf{X}}_d)\},$$

where $\langle\ ,\ \rangle$ denotes an ordered pair. We endow Γ *with a lexicographic order.* One should observe that *the set of elements in Γ which is smaller than $\langle \pi(S), d(S)\rangle$ is finite*, since $d(S) \leqslant b < \infty$ by Lemma 0.6, and $\pi(S) < \infty$ as well.

We shall argue, by way of contradiction, assuming that R is a counterexample to the theorem with a *minimal* $\langle \pi(R), d(R)\rangle (\varepsilon\Gamma)$. If $\langle \pi(R), d(R)\rangle = \langle 1, *\rangle$, then $1 = \pi(R) = \text{p.i.d.}(R/P_1) = \cdots = \text{p.i.d.}(R/P_t)$, consequently $R/N(R)$ is commutative and by Corollary 0.5, $N(R)$ is nilpotent.

The following reduction is used very often. Let $I \subset N(R)$ be a two-sided nilpotent ideal. Then $\pi(R/I) = \pi(R)$, $d(R/I) = d(R)$. Also, $N(R)$ is nilpotent iff $N(R/I) = N(R)/I$ is nilpotent. Consequently we may argue on R/I instead.

Let λ be a finite sum of evaluations of $f(\bar{\mathbf{X}}_1,\ldots, \bar{\mathbf{X}}_{\pi^2}, \bar{\mathbf{Y}}_1,\ldots, \bar{\mathbf{Y}}_s)$ on R such that $\lambda \notin P_1 \cup \cdots \cup P_m$. This is possible since $\pi \equiv \pi(R) = \text{p.i.d.}(R/P_i)$, for $i = 1,\ldots, m$. Indeed, by induction on m we pick $\mu \notin P_2 \cup \cdots \cup P_m$, where μ is a finite sum of evaluations of $f(\bar{\mathbf{X}}_1,..., \bar{\mathbf{X}}_{\pi^2}, \bar{\mathbf{Y}}_1,..., \bar{\mathbf{Y}}_s)$. Then, if $\mu \notin P_1$ pick $\lambda = \mu$. Otherwise, let $\lambda_1 \notin P_1$, $\lambda_1 \in P_2 \cap \cdots \cap P_m$ be an evaluation of $f(\bar{\mathbf{X}}_1,..., \bar{\mathbf{Y}}_s)$ (evaluating $f(\bar{\mathbf{X}}_1,..., \bar{\mathbf{Y}}_s)$ on R/P_1 and using $\bigcap_{i\neq 1} P_i \not\subset P_1$). Now $\lambda = \lambda_1 + \mu$ will do. Also, $\lambda \in P_{m+1} \cap \cdots \cap P_t$ (if $m \lneqq t$) since p.i.d.$(R/P_i) \lneqq \pi(R)$ for $i = m+1,..., t$, and $[\lambda, R] \subset P_j$, $j = 1,..., m$. Consequently, $[x_i, \lambda] \in N(R) = \bigcap_{j=1}^t P_j$, $i = 1,..., k$. Let $I = \sum_{i=1}^k R[x_i, \lambda]R$, then I is nilpotent (Theorem 0.1) and in R/I the canonical image of λ is *central.* Following our previous reduction we may therefore assume that λ is central in R.

Let $R_0 = R/\lambda R$. If $\lambda R = R$, then $N(R) \subseteq \lambda R$. So, suppose $R_0 \neq \{0\}$. We claim that $\langle \pi(R_0), d(R_0)\rangle \lneqq \langle \pi(R), d(R)\rangle$. Indeed let $P \supset \lambda R$ be a minimal prime ideal containing λR such that p.i.d.$(R/P) =$ p.i.d.$(R_0/P_0) = \pi(R_0)$ and k.d.$(R/P) =$ k.d.$(R_0/P_0) = d(R_0)$, where $P_0 \equiv P/\lambda R$. Obviously, p.i.d.$(R/P) \leqslant \pi(R)$. If p.i.d.$(R/P) \lneqq \pi(R)$, then $\pi(R_0) =$ p.i.d.$(R/P) \lneqq \pi(R)$ and consequently, $\langle \pi(R_0), d(R_0)\rangle \lneqq \langle \pi(R), d(R)\rangle$. Suppose therefore that p.i.d.$(R/P) = \pi(R) =$ p.i.d.(R/P_i), $i = 1,..., m$. Now $P \supseteq N(R) = \bigcap_{j=1}^t P_j$, yields $P \supset P_s$ for some s. Also, p.i.d.$(R/P) = \pi(R) \geqslant$ p.i.d.$(R/P_s) \geqslant$ p.i.d.(R/P), thus p.i.d.$(R/P) =$ p.i.d.(R/P_s), that is, $s \in \{1,..., m\}$.

But $\lambda \in P \backslash P_s$ implies that $P \supsetneqq P_s$ and by Lemma 0.6 we get

$$d(R_0) = \text{k.d.}\left(\frac{R}{P}\right) \lneqq \text{k.d.}\left(\frac{P}{P_s}\right) \leqslant d(R)(<\infty).$$

That is,

$$\langle \pi(R_0), d(R_0)\rangle \lneqq \langle \pi(R), d(R)\rangle.$$

Consequently, by the minimal choice of R we get that $(N(R) + \lambda R)/\lambda R \subset N(R_0)$ is nilpotent, hence

$$N(R)^l \subseteq \lambda R, \qquad \text{for some } l. \tag{$*$}$$

We next show that there exists a nilpotent ideal I in R, such that $(R/I)_{\bar{\lambda}}$ is Azumaya, here $\bar{\lambda}$ denotes the canonical image of λ in R/I. Switching to R/I and following our previous observations ($\pi(R/I) = \pi(R)$, $d(R/I) = d(R)$, $N(R/I) = N(R)/I$), this will permit the additional assumption that R_λ is Azumaya.

We have $\lambda \in P_{m+1} \cap \cdots \cap P_t$ and therefore $P_{m+1}[1/\lambda] = \cdots P_t[1/\lambda] = R_\lambda$. Consequently, $N(R_\lambda) = P_1[1/\lambda] \cap \cdots \cap P_m[1/\lambda]$, implying that $R_\lambda/N(R_\lambda) \cong (R/N)[1/\tilde{\lambda}]$, where $\tilde{\lambda}$ is the canonical image of λ in R/N.

By the Artin–Procesi theorem [10, p. 163], we have that $(R/N)[1/\tilde{\lambda}] = R_\lambda/N(R_\lambda)$ is an Azumaya algebra of constant rank π^2 over its center ($\pi \equiv \pi(R)$). Let $\bar{u}_1,..., \bar{u}_g$ be the generators of $R_\lambda/N(R_\lambda)$ over its center. Then, $\bar{u}_i\bar{u}_j = \sum_h \bar{\alpha}_{ijh}\bar{u}_h$, $i, j, h = 1,..., g$, $\bar{x}_l = \sum_h \bar{\beta}_{lh}$, $\bar{u}_h$, $h = 1,..., g$,

$l=1,\dots,k$, where $\bar{a}_{ijh}, \bar{\beta}_{lh} \in Z(R_\lambda/N(R_\lambda))$. There exists an e such that $\lambda^e u_i$, $\lambda^e a_{ijh}$, $\lambda^e \beta_{lh} \in R$, $i,j,h=1,\dots,g$, $l=1,\dots,k$, and such that $\lambda^e(u_i u_j - \sum_h a_{ijh} u_h)$, $\lambda^e(x_l - \sum_h \beta_{lh} u_h)$, $\lambda^e[\alpha_{ijh}, x_l]$, $\lambda^e[\beta_{rh}, x_l]$ are in $N(R)$ for $i,j,h=1,\dots,g$, $r,l=1,\dots,k$.

Let $u_i' = \lambda\ u_i$, $\alpha_{ijh}' \equiv \lambda^e \alpha_{ijh}$, $\beta_{rh}' = \lambda^e \beta_{rh}$. Then using the previous observations we have:

$$u_i' u_j' - \sum_h \alpha_{ijh}' u_h',\ \lambda^{2e} x_l - \sum_h \beta_{lh}' u_h',\ [\alpha_{ijh}', x_l],\ [\beta_{rh}', x_l]$$

$$\text{are in } N(R) \text{ for } i,j,h=1,\dots,g,\ r,\ l=1,\dots,k. \tag{2}$$

Let $f(\bar{\mathbf{X}}_1,\dots,\bar{\mathbf{X}}_{\pi^2}, \bar{\mathbf{Y}}_1,\dots,\bar{\mathbf{Y}}_s)$ be the polynomial appearing in Theorem 0.7; then $f(a_1,\dots,a_{\pi^2+s}) \in P_{m+1} \cap \cdots \cap P_t$ for every $a_1,\dots,a_{\pi^2+s} \in R$, since p.i.d.$(R/P_i) < \pi \equiv \pi(R)$ for $i=m+1,\dots,t$. Also, $[f(a_1,\dots,a_{\pi^2+s}), x_l] \in P_1 \cap \cdots \cap P_m$ for $l=1,\dots,k$, $a_1,\dots,a_{\pi^2+s} \in R$, since $f(\bar{\mathbf{X}}_1,\dots,\bar{\mathbf{X}}_{\pi^2}, \bar{\mathbf{Y}}_1,\dots,\bar{\mathbf{Y}}_s)$ is a central polynomial on $R/P_1 \cap \cdots \cap P_m$. Consequently,

$$[f(u_{i_1}',\dots,u_{i_{\pi^2+s}}'), x_l] \in N(R) \quad \text{for} \quad i_1,\dots,i_{\pi^2+s} \in \{1,\dots,g\} \quad \text{and} \quad l=1,\dots,k. \tag{3}$$

Let L be the two-sided ideal generated by the finite number of elements appearing in (2) and (3). Then, L is a *finitely generated* two-sided ideal in R, $L \subset N(R)$, and by Theorem 0.1 L is nilpotent.

One further verifies that $\pi(R/L) = \pi(R)$, $d(R) = d(R/L)$, so we may argue on $R/L \equiv \tilde{R}$. Now (2) implies that $\tilde{R}[1/\tilde{\lambda}]$ is a *finitely generated module* over its center, with generators $\tilde{u}_1',\dots,\tilde{u}_g'$; where $\tilde{X}$ denotes the canonical image of $X \in R$ in $\tilde{R}$. Combining this last fact with (3) and the multilinearity of $f(\mathbf{X}_1,\dots,\mathbf{Y}_s)$ we get that $\tilde{R}[1/\tilde{\lambda}]$ satisfies the identity $[f(\bar{\mathbf{X}}_1,\dots,\bar{\mathbf{X}}_{\pi^2}, \bar{\mathbf{Y}}_1,\dots,\bar{\mathbf{Y}}_s), \bar{\mathbf{X}}] = 0$.

Now, Theorem 0.7 implies that $\tilde{R}[1/\tilde{\lambda}]$ is an Azumaya algebra of constant rank π^2 over its center.

Starting again with fresh notations we may assume that

$$R_\lambda \equiv R[1/\lambda], \qquad \lambda \in Z(R), \tag{4}$$

is an Azumaya algebra of constant rank $\pi(R)^2$, over its center. Now, by Proposition 2.1 there exists e such that $\lambda^e R \subseteq Zb_1 + \cdots + Zb_h$, where $Z \equiv Z(R)$, $b_i \in R$, for $i=1,\dots,h$, and $Zb_1 + \cdots + Zb_h$ is a ring.

We next show that given a *finite* number of identities of $M_\pi(\mathbb{Z})$, $g_1,\dots,g_\alpha$, we can assume that R satisfies $g_1,\dots,g_\alpha$. We prove it with $g=g_1$, the extension to $\alpha > 1$ is obvious. We may assume that $g \equiv g(X_1,\dots,X_s)$ is homogeneous of degree a.

In the process of linearization of g we obtain a finite number of polynomials $g=g^{(1)}, g^{(2)},\dots,g^{(t)}$, where $g^{(i)}$ is an identity of $M_\pi(\mathbb{Z})$ for

$i=1,\ldots,t$. Let $r_1,\ldots,r_s \in R$; then, by Proposition 2.1, $\lambda^e r_i = \sum_1^h \alpha_{ij} b_j$. Consequently,

$$\lambda^{ea} g(r_1,\ldots,r_s) = g(\lambda^e r_1,\ldots,\lambda^e r_s) = g\left(\sum_j \alpha_{1j} b_j,\ldots,\sum_j \alpha_{sj} b_j\right)$$
$$= \sum_{i,j_1,\ldots,j_s} \alpha^{(i)} g^{(i)}(b_{j_1},\ldots,b_{j_s}),$$

where $\alpha^{(i)}$ is a monomial on $\{\alpha_{ij}\}$.

Now $\{g^{(i)} \mid i=1,\ldots,t\}$ are identities of $M_\pi(\mathbb{Z})$ and therefore vanish on $R[1/\lambda]$. Also, the set $\{g^{(i)}(b_{j_1},\ldots,b_{j_s}) \mid i=1,\ldots,t,\ j_l=1,\ldots,h\}$ is a finite set of elements in R, implying that there exists n such that $\lambda^n g^{(i)}(b_{j_1},\ldots,b_{j_s})=0$ for $i=1,\ldots,t$, $j_l=1,\ldots,h$. Consequently, $\lambda^{ea+n} g(r_1,\ldots,r_s)=0$ and we have that $\lambda^{ea+n} V(g)=\{0\}$, where $V(g)$ is the ideal in R generated by all the evaluations of g. Now *if* $N(R/V(g))$ is nilpotent of index p, then $N(R)^p \subset V(g)$, and therefore $\lambda^{ea+n} N(R)^p = \{0\}$. Now (*) implies that $N(R)^{(ea+n)l+p}=\{0\}$.

Taking $\alpha=2$, $g_1=g$, $g_2=f_{\pi^2}$ as in Theorem 1.2. Then, the previous reasoning shows that we may assume that R satisfies the identities g, f_{π^2}. Consequently, by Corollary 1.3, $V_c^p(R)\, V(d_{\pi^2})=\{0\}$, for some p and c.

Recall that $V_c(R)$ denotes the ideal in R generated by all the evaluations of all identities of $M_c(\mathbb{Z})$ for some c. The required result now follows, since $N(R/V_c(R))$ is nilpotent of index q for some q by [15, Theorem 2], hence $N(R)^q \subseteq V_c(R)$. Also, $\pi(R/V(d_{\pi^2})) \leqslant \pi - 1 < \pi(R) \equiv \pi$ and therefore by minimality $N(R/V(d_{\pi^2}))$ is nilpotent of index t, implying that $N(R)^t \subseteq V(d_{\pi^2})$. Consequently,

$$N(R)^{qp+t} \subseteq V_c^p(R) \cdot V(d_{\pi^2}) = \{0\}. \qquad \text{Q.E.D.}$$

As an immediate corollary we have the following:

THEOREM 2.3. *Let $R=F\{x_1,\ldots,x_k\}$ be a* p.i. *ring, where F is a central subfield. Then $N(R)$ is nilpotent.*

We also have the following important corollary.

THEOREM 2.4. *Let $R=F\{x_1,\ldots,x_k\}$ be a* p.i. *ring, F a central subfield. Then, R satisfies all the identities of $M_n(\mathbb{Z})$ for some n. Equivalently, R is an homomorphic image of the ring of $n\times n$ generic matrices $F\{\bar{\xi}_1,\ldots,\bar{\xi}_k\}$.*

Proof. We apply [9, Theorem 10]. Indeed, if F is infinite, then this is exactly a consequence of [9, Theorem 10]. If F is finite, let $\bar{F}$ be its algebraic closure; then $R \subset R_{\bar{F}} \equiv R \otimes \bar{F}$, $R_{\bar{F}}$ a p.i. ring, and by Theorem 2.3 its nil radical is nilpotent. Now, by [9, Theorem 10] $R_{\bar{F}}$ satisfies all the identities of $M_n(\mathbb{Z})$ for some n ($\bar{F}$ is infinite) and consequently, so does R.

Before beginning the proof of Theorem 2.5, we should point out the difficulty in proving the general statement. Following the notations of Theorem 2.2, say P is a minimal prime containing λ and p.i. $d(R/P) = \pi(R)$. Suppose that $d(R) = \infty$; then it may well happen that

$$d(R_0) = \text{k.d.}\left(\frac{R}{P}\right) = \text{k.d.}\left(\frac{R}{P_s}\right) = d(R), \qquad \text{although } P \supsetneqq P_s.$$

So the obstacles are cases where $d(R) = \infty$. This difficulty is avoided if k.d.$(\Lambda) < \infty$ (or k.d.$(R) < \infty$), which is the case of Theorem 2.2.

THEOREM 2.5. *Let* $R = \Lambda\{x_1,..., x_k\}$ *be a* p.i. *ring, satisfying a monic polynomial identity, and* Λ *is a central noetherian ring.*

Then, $N(R)$*, the nil radical of* R*, is nilpotent.*

Proof. Using the same notations as in Theorem 2.2, we shall prove by induction on $\pi(R)$ that there exists $\alpha_1,..., \alpha_q$, evaluations of $f(\bar{\mathbf{X}}_1,..., \bar{\mathbf{Y}}_s)$ such that

$$N(R)^l \subseteq R\alpha_1 R + \cdots + R\alpha_q R, \qquad \text{for some } l, q. \qquad (*)$$

Let $G = \{S \mid S = \Lambda\{y_1,..., y_k\},\ 1 \in S,\ S \text{ satisfies the identity } P(\bar{\mathbf{X}}_1,..., \bar{\mathbf{X}}_d)\}$.

We assume the truth of $(*)$ for every $S \in G$ with $\pi(S) \lneqq \pi(R)$.

The case $\pi(R) = 1$ is easily checked as in Theorem 1.1.

As in Theorem 2.2, let $P_1,..., P_m$ be the minimal prime ideals of R, satisfying p.i.d.$(R/P_i) = \pi(R)$, $i = 1,..., m$. Let $\alpha \notin P_1 \cup \cdots \cup P_m$, be finite sum of evaluations of $f(X_1,..., Y_s)$ (again, as Theorem 2.2), and let $R_0 = R/R\alpha R$. Obviously $\pi(R_0) \leqslant \pi(R)$.

If $\pi(R_0) \lneqq \pi(R)$, we get by induction that

$$N(R_0)^l \subseteq R\tilde{\beta}_1 R + \cdots + R_0\tilde{\beta}_r R_0,$$

where $\tilde{\beta}_i$ is the image in R_0 of the evaluation β_i of $f(\bar{\mathbf{X}}_1,..., \bar{\mathbf{Y}}_s)$.

Now $(N(R) + R\alpha R)/R\alpha R \subseteq N(R_0)$ and therefore

$$N(R)^l \subseteq R\beta_1 R + \cdots + R\beta_r R + R\alpha R.$$

So, we may assume that $\pi(R_0) = \pi(R) \equiv \pi$. We use now a second induction in order to prove $(*)$. More precisely, we prove for each $S \in G$ with $\pi(S) = \pi$, $d(S) < \infty$ that $(*)$ is valid, by induction on $d(S)$.

Let $d(R) = 0$ and $\alpha \in R_0$ be as above. Let P be a minimal prime ideal above α; then since $P \supseteq N(R) = \bigcap_{i=1}^t P_i$, $P \supset P_s$ for some $s \in 1,..., t$. Also $\alpha \in P \backslash P_s$ implies that $P \supsetneqq P_s$. Now, since $d(R) = 0$ we have that k.d.$(R/P_i)=0$, $i=1,\ldots,m$, and therefore $s \notin \{1,\ldots,m\}$. Consequently, p.i.d.$(R/P) <$ p.i.d.$(R/P_s) \lneqq \pi(R)$. Hence $\max\{\text{p.i.d.}(R/P) \mid \alpha \in P,\ P$ is prime$\} = \pi(R_0) \lneqq \pi(R)$. So, by the first induction applied to R_0, $N(R_0)^l \subseteq \sum_{i=1}^r R_0\bar{\beta}_i R_0$, where $\{\beta_i\}$ are evaluations of $f(\bar{\mathbf{X}}_1,\ldots,\bar{\mathbf{Y}}_s)$. Thus

$$N(R)^l \subseteq R\alpha R + R\beta_i R + \cdots + R\beta_r R.$$

Suppose we proved it for $S \in G$ with $d(S) \lneqq d(R)$. Let α be as before and $R_0 = R/R\alpha R$. If $\pi(R_0) \lneqq \pi(R)$, then we are done by the first induction. So suppose $\pi(R_0) = \pi(R) \equiv \pi$. We claim that $d(R_0) \lneqq d(R)$. For let $P_1^{(0)},\ldots,P_{m_0}^{(0)}$ be the minimal prime ideals above $R\alpha R$, satisfying $\pi(R/P_i^{(0)}) = \pi(R_0)$, and k.d.$(R/P_i^{(0)}) = d(R_0)$, $i=1,\ldots,m_0$.

Given $i \in \{1,\ldots,m_0\}$, there exists $s \in \{1,\ldots,m\}$ such that $P_i^{(0)} \supset P_s$. Indeed $P_i^{(0)} \supseteq N(R) = \bigcap^t P_j$ implies $P_i^{(0)} \supset P_s$ for some s. Now,

$$\text{p.i.d.}(R/P_i^{(0)}) = \pi(R_0) = \pi(R) \geqslant \text{p.i.d.}(R/P_s) \geqslant \text{p.i.d.}(R/P_i^{(0)})$$

(the last inequality is valid due to $P_i^{(0)} \supseteq P_s$). So, $\pi = \text{p.i.d.}(R/P_i^{(0)}) =$ p.i.d.(R/P_s) and $s \in \{1,\ldots,m\}$. Thus, $d(R_0) = \text{k.d.}(R/P_i^{(0)}) \leqslant \text{k.d.}(R/P_s) \leqslant d(R) \lneqq \infty$, also $\alpha \in P_i^{(0)} \backslash P_s$, $P_i^{(0)} \supsetneqq P_s$ and since k.d.$(R/P_s) \lneqq \infty$ we have $d(R_0) = \text{k.d.}(R/P_i^{(0)}) \lneqq d(R)$. Consequently, by the second induction $N(R_0)^l \subseteq R_0\tilde{\beta}_1 R_0 + \cdots + R_0\tilde{\beta}_r R_0$ or $N(R)^l \subseteq R\beta_1 R + \cdots + R\beta_r R + R\alpha R$, where β_i are evaluations of $f(\bar{\mathbf{X}}_1,\ldots,\bar{\mathbf{Y}}_s)$.

The remaining case is to show the validity of $(*)$ for $S \in G$ with $\pi(S) = \pi$ and $d(S) = \infty$. So we assume that $d(R) = \infty$.

Let α, R_0, $P_1^{(0)},\ldots,P_{m_0}^{(0)}$ be as before. We have $\pi(R_0) = \pi(R) = \pi$. Also, as before, for each $i \in \{1,\ldots,m_0\}$ there exists $j \in \{1,\ldots,m\}$ such that $P_i^{(0)} \supsetneqq P_j$, p.i.d.$(R/P_i^{(0)}) =$ p.i.d.$(R/P_j) = \pi$ and k.d.$(R/P_i^{(0)}) \leqslant$ k.d.$(R/P_j) \leqslant d(R)$. If we have k.d.$(R/P_i^{(0)}) \lneqq$ k.d.(R/P_j) then $d(R_0) = \text{k.d.}(R/P_i^{(0)}) \lneqq d(R) = \infty$ and we apply the second induction on R_0 and finish as before. So, assuming that $d(R_0) = \infty$, we have $\alpha \in \{\bigcap_{i=1}^{m_0} P_i^{(0)}\} \backslash \{\bigcap_{i=1}^m P_i\}$, $P_i^{(0)} \supset P_j$, for some $j = j(i)$, $j \in \{1,\ldots,m\}$, $i=1,\ldots,m_0$, hence

$$\left\{\bigcap_{i=1}^{m_0} P_i^{(0)}\right\} \supsetneqq \left\{\bigcap_{i=1}^{m} P_i\right\}.$$

Continuing the process with R_0, let γ be an evaluation of $f(\bar{\mathbf{X}}_1,\ldots,\bar{\mathbf{X}}_{\pi^2},\bar{\mathbf{Y}}_1,\ldots,\bar{\mathbf{Y}}_s)$ satisfying $\gamma \notin P_1^{(0)} \cup \cdots \cup P_{m_0}^{(0)}$ and let $R_1 = R/(R\gamma R + R\alpha R)$. Let $P_1^{(1)},\ldots,P_{m_1}^{(1)}$ be the minimal prime ideal above $R\gamma R + R\alpha R$ such that

$$\pi(R_1) = \text{p.i.d.}\left(\frac{R}{P_i^{(1)}}\right), \qquad d(R_1) = \text{k.d.}\left(\frac{R}{P_i^{(1)}}\right), \qquad i=1,\ldots,m_1.$$

Let $i \in \{1,..., m_1\}$ we want to show the existence of $s \in \{1,..., m_0\}$ $s = s(i)$ such that $P_i^{(1)} \supset P_s^{(0)}$. Indeed, $P_i^{(1)} \supset R\gamma R + R\alpha R \supset R\alpha R$, hence $P_i^{(1)} \supseteq P$, where P is a minimal prime above $R\alpha R$. If p.i.d.$(R/P) \lneqq \pi(R_0)$ then $\pi(R_1) =$ p.i.d.$(R/P_i^{(1)}) \leqslant$ p.i.d.$(R/P) \lneqq \pi(R_0) = \infty$ and we use the first induction on R_1 to show that $N(R)^l \subseteq R\beta_1 R + \cdots + R\beta_r R + R\gamma R + R\alpha R$, for some l, where $\beta_1,..., \beta_r$ are evaluations of $f(\overline{\mathbf{X}}_1,..., \overline{\mathbf{Y}}_s)$. So $\pi(R_1) = \pi(R_0) = \pi$. If, in

$$d(R_1) = \text{k.d.} \left(\frac{R}{P_i^{(1)}}\right) \leqslant \text{k.d.} \left(\frac{R}{P}\right) \leqslant d(R_0) = \infty,$$

one of the inequalities is *strict*, then $d(R_1) \lneqq \infty$ and we use the second induction on R_1 to finish. Thus, $\infty = d(R_1) = \text{k.d.}(R/P)$ and therefore $P \in \{P_1^{(1)},..., P_{m_1}^{(1)}\}$, or equivalently, $P = P_s^{(1)}$, for some $s \in \{1,..., m_1\}$. As before, we have

$$\left\{\bigcap_{i=1}^{m_1} P_i^{(1)}\right\} \supsetneqq \left\{\bigcap_{i=1}^{m_0} P_i^{(0)}\right\} \qquad \text{since} \quad \gamma \in \left\{\bigcap_{i=1}^{m_1} P_i^{(1)}\right\} \Big\backslash \left\{\bigcap_{i=1}^{m_0} P_i^{(0)}\right\}.$$

Continue the process if needed, we get an ascending chain of semiprime ideals

$$\left\{\bigcap_{i=1}^{m} P_i\right\} \subsetneqq \left\{\bigcap_{i=1}^{m_0} P_i^{(0)}\right\} \subsetneqq \left\{\bigcap_{i=1}^{m_1} P_i^{(1)}\right\} \subset \cdots \subset \left\{\bigcap_{i=1}^{m_g} P_i^{(g)}\right\} \subset \cdots,$$

which must terminate at some stage g by [10, p. 106, Corollary 2.2]. Hence for some g either $\pi(R_g) \lneqq \pi$ or $d(R_g) \lneqq \infty$ and applying the first or second induction respectively on R_g we get

$$N(R)^l \subseteq R\alpha_1 R + \cdots + R\alpha_r R, \qquad \text{for some } l, r. \tag{$*$}$$

Now, $\sum_{i=1}^r R[R, \alpha_i]R = \sum_{i=1}^r \sum_{j=1}^k R[x_j, \alpha_i]R$, $=I$ is a finitely generated ideal which is contained in $N(R)$, so I is nilpotent (Theorem 0.1), and switching to R/I we may assume that $\alpha_1,..., \alpha_r$ are central in R.

Our next step is to show that we may assume that R_{α_i} is Azumaya for $i = 1,..., r$. This is carried exactly as in Theorem 1.1 for R_λ, and we get

$$\begin{gathered} \alpha_m^{e_m} R \subseteq Zb_1^{(m)} + \cdots + Zb_{h_m}^{(m)}, \qquad \text{where} \quad Z \equiv Z(R) \quad \text{and} \\ Zb_1^{(m)} + \cdots + Zb_{h_m}^{(m)} \text{ is a subring of } R \text{ for } m = 1,..., r. \end{gathered} \tag{$**$}$$

Now following the argument in Theorem 2.2, we get for each α_m separately, that $N(R)^{p_m} \alpha_m^{e_m g_m} = \{0\}$, $m = 1,..., r$, and if we pick

$$p = \max\{p_m \mid m = 1,..., r\}, \qquad g = \max\{g_m \mid m = 1,..., r\},$$
$$e = \max\{e_m \mid m = 1,..., r\},$$

we get $N(R)^p \alpha_m^{eg} = \{0\}$, $m = 1,..., r$.

Now by $(*)$, $N(R)^l \subseteq \sum_{m=1}^r \alpha_m R$, hence $N(R)^{lreg} \subseteq \sum_{m=1}^r \alpha_m^{eg} R$ and consequently $N(R)^{lreg+p} = \{0\}$.

Appendix

We reproduce here the results of Shirshov and Latyshev, which are used in Theorem 0.1. We first quote Shirshov's theorem and then show how this implies Latyshev's result. (We follow [8, Proposition 12].) We then reprove Shirshov's theorem.

THEOREM A.1 [15, Theorem 3]. *Suppose $A = \Lambda\{y_1,..., y_t\}$ is an associative ring satisfying a monic (multilinear identity of degree d, where $\Lambda \subseteq Z(A)$.*

Then, there exist a finite number of words $v_1,..., v_s$ in the generators $y_1,..., y_t$, length $(v_i) \leqslant d$ (with respect to $y_1,..., y_t$), and a number N such that any word w in $y_1,..., y_t$ with length $(w) \geqslant N$ has the following representation in A.

$$w = \sum \lambda_v v_{i_1}^{l_{i_1}} \cdots v_{i_h}^{l_{i_h}}, \qquad \text{where} \quad \lambda_v \in \Lambda,\ l_{i_j} \geqslant 0, \text{ and } h \text{ is fixed.}$$

Moreover the decomposition of each monomial $v_{i_1}^{l_{i_1}} \cdots v_{i_h}^{l_{i_h}}$ i with respect to $y_1,..., y_t$ is a permutation of w with respect to $y_1,..., y_t$.

THEOREM A.2 [8, Proposition 12]. *Let $A = \Lambda\{a_1,..., a_k\}$ be an associative ring satisfying a monic multilinear identity of degree d, where $\Lambda \subseteq Z(A)$. Let $I \subset N(R)$ be a two-sided ideal and*

$$I = \sum_{i=1}^{m} Rn_i R.$$

Then, I is nilpotent.

Proof. We regard A as being generated by $a_1,..., a_k$, $n_1,..., n_m$ and let $v_1,..., v_s$, N be as in Theorem A.1 (with respect to these generators). Recall that length $(v_i) \leqslant d$. We observe that some of the v_i's are in I.

Let M be the largest index of nilpotency of such v_i's. Let $q = hdM + 1$ and $p = \max(N, q)$. We shall show that $I^p = \{0\}$. Let $w \in I^p$; then w is a monomial in $a_1,..., a_k$, $n_1,..., n_m$ and by Theorem A.1 $w = v_{i_1}^{l_{i_1}} \cdots v_{i_h}^{l_{i_h}}$ (length $(w) \geqslant p \geqslant N$). We have that p elements of $n_1,..., n_m$ (with repetitions) appear in w and therefore there exists a j such that $v_{i_j}^{l_{i_j}}$ contains $dM + 1$ elements of $n_1,..., n_m$ (with repetitions) since $p \geqslant sdM + 1$. Thus length $(v_{i_j}^{l_{i_j}}) \geqslant dM + 1$. Also, since length $(v_{i_j}) \leqslant d$ we have that length $(v_{i_j}^{l_{i_j}}) \leqslant dl_{i_j}$. Consequently

$l^{i_j} \geqslant M$. Now since $v_{i_j} \in I$ we have that $v_{i_j}^M = 0$ and therefore $v_{i_j}^{l_{i_j}} = 0$, implying that $w = 0$. Q.E.D.

We next reproduce the proof of Theorem A.1, mainly because there is no available reference in English. In order to do so we use another Shirshov theorem [12, p. 206]. So we recall the basic definitions of the total ordering which are introduced on the free monoid generated by $X_1,..., X_k$ and 1. We specify that $X_1 < X_2 < \cdots < X_k$ and that $u < 1$ for any $u \neq 1$, and $u = X_{i_1} \cdots X_{i_r} > v = X_{j_1} \cdots X_{j_s}$ if either u is an initial segment of v ($v = uw$, $w \neq 1$) or $i_1 = j_1,..., i_k = j_k$ and $i_{k+1} > j_{k+1}$ for $k < \min(r, s)$. Also, if $w = w_1 \cdots w_m$ is a factorization of w to subwords, then w is called *dominant* if $w > w_{\sigma(1)} w_{\sigma(2)} \cdots w_{\sigma(m)}$ for every permutation $\sigma \neq 1$ in the symmetric groups Σ_m.

THEOREM [12, p. 206; 16]. *If k, m, M are positive integers, there exists a positive integer $g(k, m, M)$ such that any word in the free monoid generated by* 1, $X_1,..., X_k$ *of length* $\geqslant g(k, m, M)$ *contains a subword w of one of the following forms*:

(i) $w = u^M$, $1 < \mathit{length}(u) \leqslant m$;

(ii) *w has a dominant factorization of length m.*

We prove now Theorem A.1. We begin with:

DEFINITION. The *height* of a monomial a with respect to a set of generators $v_1,..., v_s$ is h if $a = v_{i_1}^{l_1} \cdots v_{i_h}^{l_h}$ and h is minimal.

We prove the theorem by induction on the lexicographic order described earlier (lifting the monomials to $\Lambda\{X_1,..., X_k\}$ and doing the induction there). Let $m = d$, $M = d$ and $g(k, d, d) = N$. Let $v_1,..., v_s$ be the set of the monomials in $x_1,..., x_k$ of length $\leqslant d$. Let a be a monomial in $\Lambda\{x_1,..., x_k\}$ of height $>N$. We obviously have that length $(a) \geqslant N$.

If a has a dominant subword w, $a = cwe$ $w = w_1 \cdots w_d$, then since R satisfies an identity of degree d,

$$w = \sum_{\sigma \neq 1} \alpha_\sigma w_{\sigma(1)} \cdots w_{\sigma(d)}$$

and

$$a = \sum_{\sigma \neq 1} \alpha_0 c w_{\sigma(1)} \cdots w_{\sigma(d)} e,$$

and each $cw_{\sigma(1)} \cdots w_{\sigma(d)} e$ being of smaller order has the required expression by induction. If $w = u^d$ we claim that a contains a subword of the form $b^d b'$, where length (b), length $(b') \leqslant d$ and b' is *not* an initial segment of b. Indeed

$a = cu^d e$ and by collecting initial subwords equal to u we may assume that $a = cu^{d_1}e$, $d_1 \geqslant d$ and u does not appear in e as an initial segment. Thus, either length$(e) \geqslant d$ and we can find an initial segment of e, b' length $(b') \leqslant d$ which is *not* an initial segment of u and then we take $b \equiv u$, or length $(e) < d$ and e is an initial segment of u, $u = ee_1$. Hence $a = cu^{d_1}e = ce(e_1 e)^{d_1} = ce(u')^{d_1}$, where $u' = e_1 e$. Observe that height $(a) \leqslant$ height$(ce) + 1$. Now, if height $(ce) < N$ then $(a) \leqslant N$, a contradiction. So height $(ce) \geqslant N$ implying that length $(ce) \geqslant N$, and as before, we may assume that ce has a subword of the form $(u'')^d$, length $(u'') \leqslant d$, $a = c_1(u'')^d e_2(u')^{d_1}$.

Equivalently $a = c_1(u'')^{d_2} e_2^1$, where in e_2^1, u'' does not appear as an initial segment, $d_2 \geqslant d$. As before, either e_2^1 has an initial segment b^1, length $(b^1) \leqslant d$ and b^1 is *not* an initial segment of u'' and we take $b = u''$, or e_2^1 is an initial segment of u'', length $(e_2^1) < d$. So, $u'' = e_2^1 e_3$ and $a = c_1 e_2^1 (u''')^{d_2}$, where $u''' \equiv e_3 e_2^1$. Now

$$\text{length}(u''')^{d_2} = \text{length}\{(u'')^d (u')^{d_1}\} - \text{length } e_2^1 + \text{length } e_2$$
$$\geqslant \text{length } (u')^{d_1} + 1, \text{ since length } (e_2^1) < d.$$

Consequently, length $(c_1 e_2^1) <$ length(ce), and by continuing the process with $c_1 e_2^1$ we must stop after a finite step bounded by length (a).

Thus, a must contain a subword of the form $b^d b^1$ with length (b), length $(b^1) \leqslant d$ and b' *is not an initial segment of* b. Also, if b is an initial segment of b^1 we can write $b^d b^1 = b^{d'} b''$, $d' \geqslant d$ and b'' is *not* an initial segment of b (if $b^d b' = b^{d''}$ then length$(b) < d/2$ and we take $b^1 \equiv b^2$).

Now the number of choices for pairs b, b' is finite ($\leqslant s^2$), and by taking a word a of height bigger then $Ns^2 d$, there exists a subword of the form

$$b^d b' \alpha_1 b^d b' \alpha_2 \cdots b^d b' \alpha_{d-1} b^d b' \alpha_d,$$

where b, b' are as above.

If $b > b'$ the previous subword can be written in the form

$$(b^d b' \alpha_1 b)(b^{d-1} b' \alpha_2 b^2)(b^{d-2} b' \alpha_3 b^3) \cdots (bb' \alpha_d),$$

and this word is obviously dominant. Thus we write it via the identity as a sum of monomials of smaller order and the required result follows by the induction on the order.

If $b' > b$, we can write the word in the form

$$b^d (b' \alpha_1 b^{d-1})(bb' \alpha_2 b^{d-2}) \cdots (b^{d-1} b' \alpha_d)$$

and we get again a dominant word of degree d *since* $b' > b$ *and* b' *is not an initial segment of* b, again the previous argument applies. Q.E.D.

ACKNOWLEDGMENTS

I would like to thank O. Gabber and the referee for pointing out errors in earlier versions of this paper.

REFERENCES

1. S. A. AMITSUR, Alternating identities, *in* "Proceedings, Ohio Univ. Ring Theory Conf.," pp. 1–14, Dekker, New York.
2. S. A. AMITSUR AND L. W. SMALL, Prime ideals in *PI*-rings, *J. Algebra* **62** (1980), 358–383.
3. A. BRAUN, Affine P.I.-rings and their generalizations, *J. Algebra* **58** (1979), 481–494.
4. A. BRAUN, On M. Artin's theorem and Azumaya algebras, *J. Algebra*, **77** (1982), 323–332.
5. G. CAUCHON, Anneaux premiers, noetheriens, à identités polynômiales, *Bull. Soc. Math. France* **104** (1976), 99–111.
6. F. DEMEYER AND E. INGRAHAM, "Separable Algebras over Commutative Rings," Lecture Notes in Mathematics No. 181, Springer-Verlag, Berlin/New York, 1971.
7. A. T. KEMER, Capelli identities and nilpotency of the radical of a finitely generated PI-algebra, *Dokl. Akad. Nauk SSSR* **255** (1980), 793–797 (Russian).
8. V. N. LATYSHEV, On the complexity of non-matrix varieties of associative algebras, *Algebra i Logika* **16** (1977), 98–122.
9. J. LEWIN, A matrix representation for associative algebras, *Trans. Amer. Math. Soc.* **188** (1974), 292–308.
10. C. PROCESI, "Rings with Polynomial Identities," Dekker, New York, 1973.
11. Y. P. RAZMYSLOV, The Jacobson radical in P.I.-algebras, *Algebra and Logic* **13** (1974), 192–204.
12. L. W. ROWEN, "Polynomial Identities in Ring Theory," Academic Press, New York/London, 1980.
13. W. SCHELTER, Integral extension of rings satisfying a polynomial identity, *J. Algebra* **40** (1976), 245–257.
14. W. SCHELTER, Azumaya algebras and M. Artin's theorem, *J. Algebra* **46** (1977), 303–304.
15. W. SCHELTER, Non commutative affine P.I.-rings are catenary, *J. Algebra* **51** (1978), 12–18.
16. A. I. SHIRSHOV, Rings with identical relations, *Math. Sb.* **43** (1957), 277–283.